```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

```
df=pd.read_csv("/content/onlinefoods.csv")
```

```
df.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 388 entries, 0 to 387
Data columns (total 13 columns):
 #   Column                       Non-Null Count  Dtype  
---  ------                       --------------  -----  
 0   Age                          388 non-null    int64  
 1   Gender                       388 non-null    object 
 2   Marital Status               388 non-null    object 
 3   Occupation                   388 non-null    object 
 4   Monthly Income               388 non-null    object 
 5   Educational Qualifications   388 non-null    object 
 6   Family size                  388 non-null    int64  
 7   latitude                     388 non-null    float64
 8   longitude                    388 non-null    float64
 9   Pin code                     388 non-null    int64  
 10  Output                       388 non-null    object 
 11  Feedback                     388 non-null    object 
 12  Unnamed: 12                  388 non-null    object 
dtypes: float64(2), int64(3), object(8)
memory usage: 39.5+ KB
```

```
df.head()
```

	Age	Gender	Marital Status	Occupation	Monthly Income	Educational Qualifications	Family size	latitude	longitu
0	20	Female	Single	Student	No Income	Post Graduate	4	12.9766	77.59
1	24	Female	Single	Student	Below Rs.10000	Graduate	3	12.9770	77.57
2	22	Male	Single	Student	Below Rs.10000	Post Graduate	3	12.9551	77.65
					No				

```
print(df.describe())
```

```
              Age  Family size    latitude   longitude       Pin code
count  388.000000   388.000000  388.000000  388.000000     388.000000
mean    24.628866     3.280928   12.972058   77.600160  560040.113402
std      2.975593     1.351025    0.044489    0.051354      31.399609
min     18.000000     1.000000   12.865200   77.484200  560001.000000
25%     23.000000     2.000000   12.936900   77.565275  560010.750000
50%     24.000000     3.000000   12.977000   77.592100  560033.500000
75%     26.000000     4.000000   12.997025   77.630900  560068.000000
max     33.000000     6.000000   13.102000   77.758200  560109.000000
```

```
print(df.isnull().sum())
```

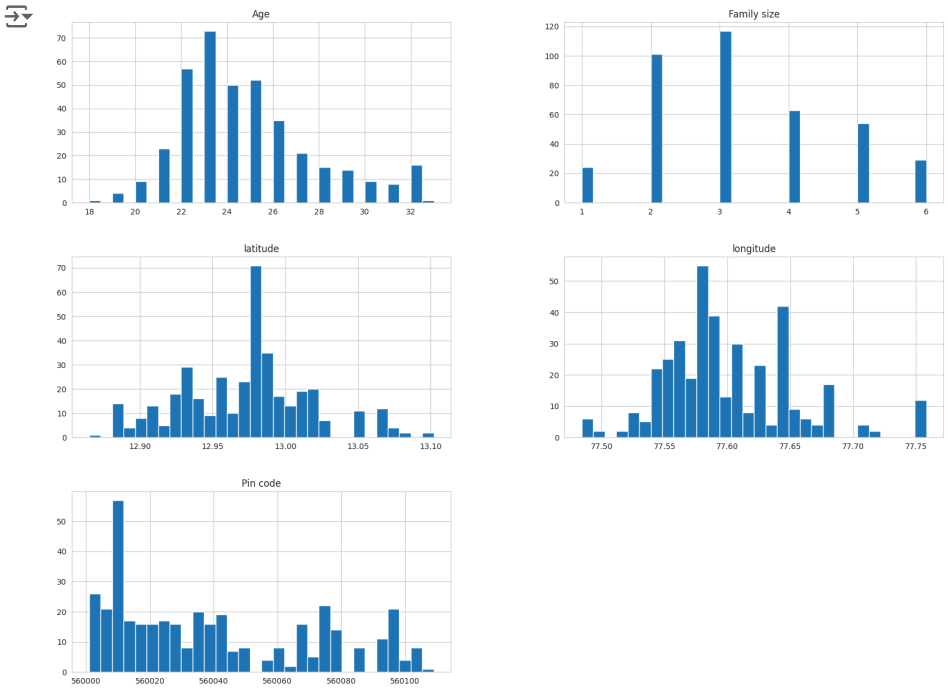
```
Age                          0
Gender                       0
Marital Status               0
Occupation                   0
Monthly Income               0
Educational Qualifications   0
Family size                  0
latitude                     0
longitude                    0
Pin code                     0
Output                       0
Feedback                     0
```

```
Unnamed: 12                0
dtype: int64
```

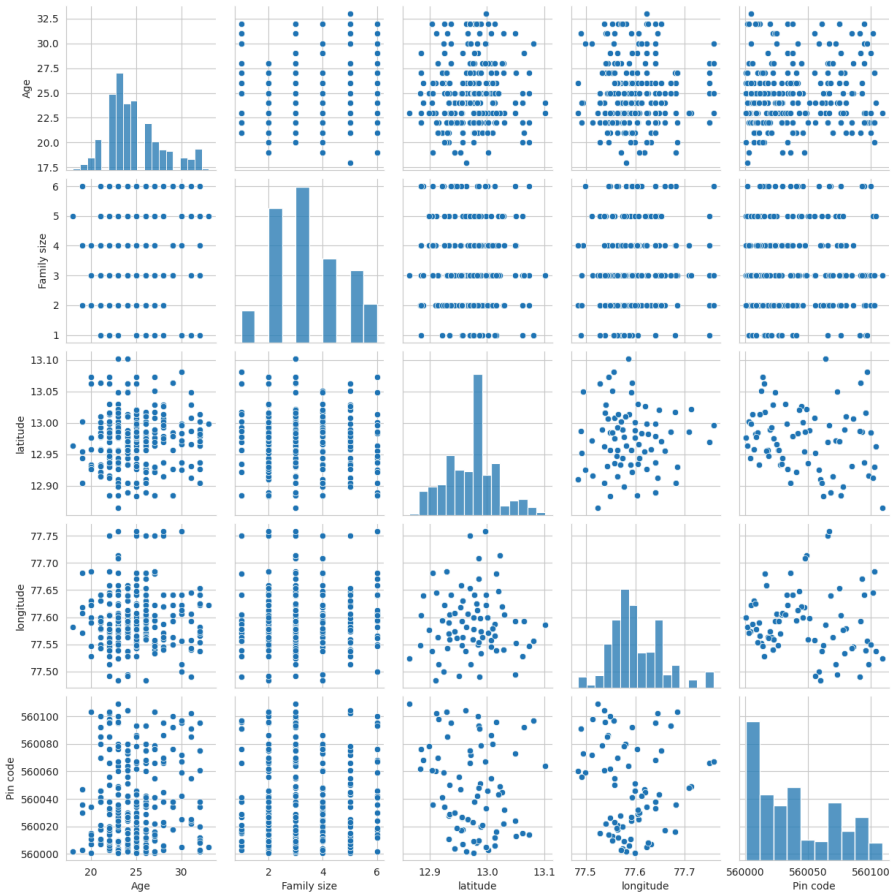
```
import matplotlib.pyplot as plt
import seaborn as sns

# Set the aesthetic style of the plots
sns.set_style("whitegrid")
```

```
df.hist(bins=30, figsize=(20, 15))
plt.show()
#Histogram
```



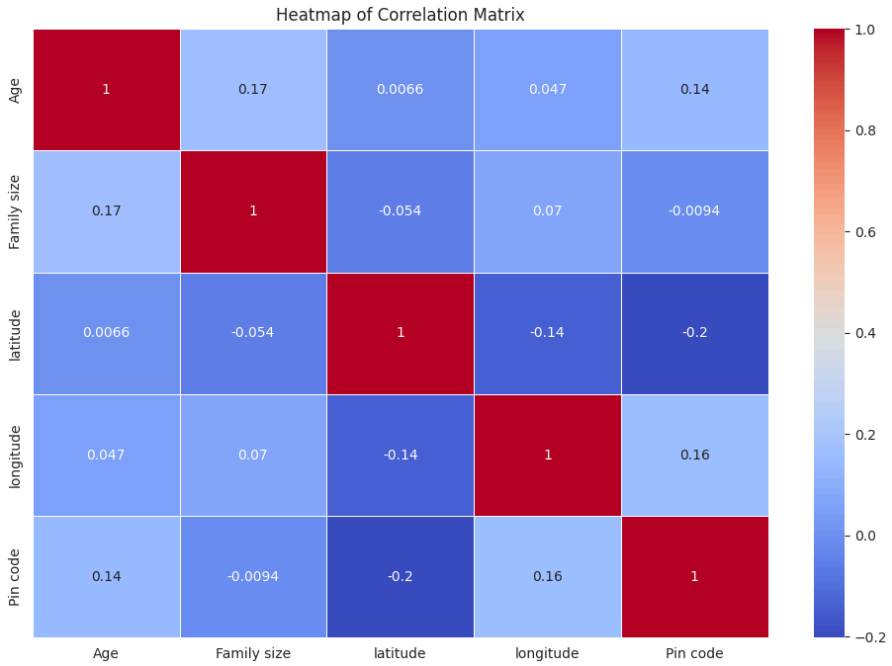

```
sns.pairplot(df)
plt.show()
#pair
```




```
numeric_df = df.select_dtypes(include=['float64', 'int64'])
```

```
numeric_df = numeric_df.dropna()
```

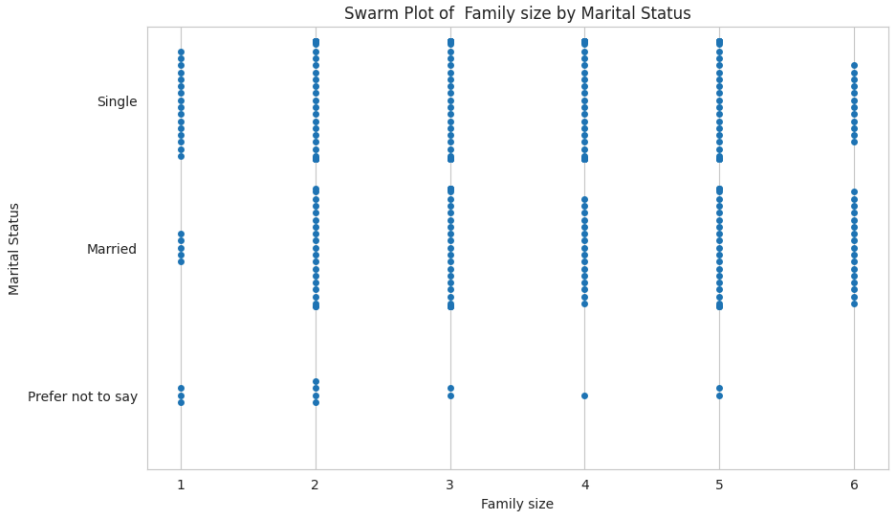
```
plt.figure(figsize=(12, 8))
correlation_matrix = numeric_df.corr()
sns.heatmap(correlation_matrix, annot=True, cmap='coolwarm', linewidths=0.5)
plt.title('Heatmap of Correlation Matrix')
plt.show()
#heat map
```



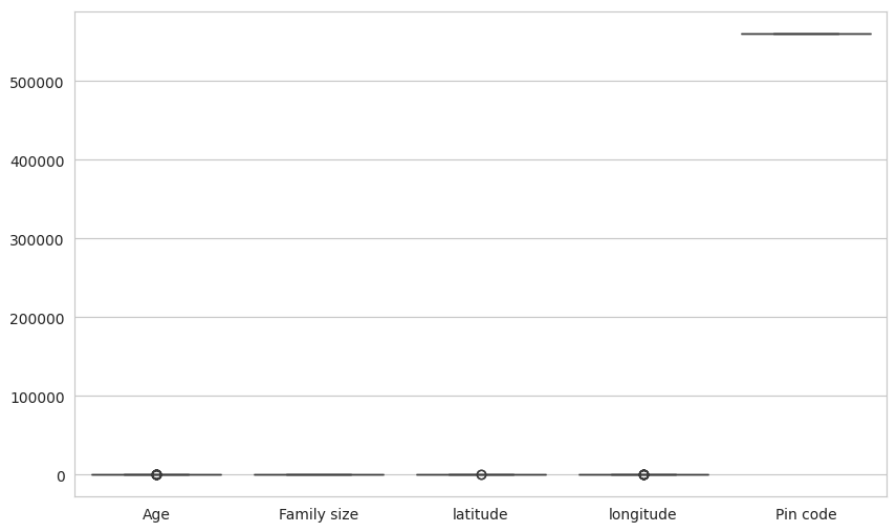

```
plt.figure(figsize=(10, 6))
sns.swarmplot(x='Family size', y='Marital Status', data=df)
plt.title('Swarm Plot of  Family size by Marital Status')


plt.show()
#swarm
```

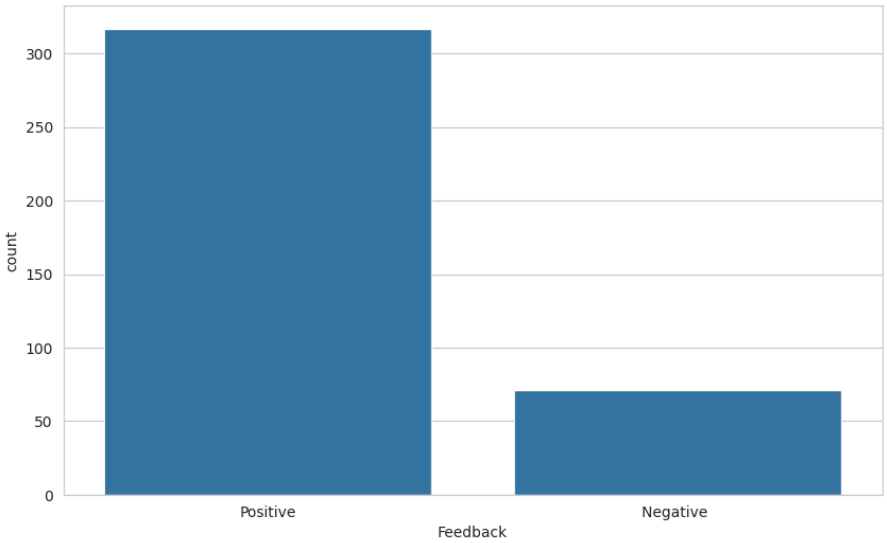
```
/usr/local/lib/python3.10/dist-packages/seaborn/categorical.py:3398: UserWarning: 55.2%
  warnings.warn(msg, UserWarning)
/usr/local/lib/python3.10/dist-packages/seaborn/categorical.py:3398: UserWarning: 64.2%
  warnings.warn(msg, UserWarning)
/usr/local/lib/python3.10/dist-packages/seaborn/categorical.py:3398: UserWarning: 17.6%
  warnings.warn(msg, UserWarning)
```



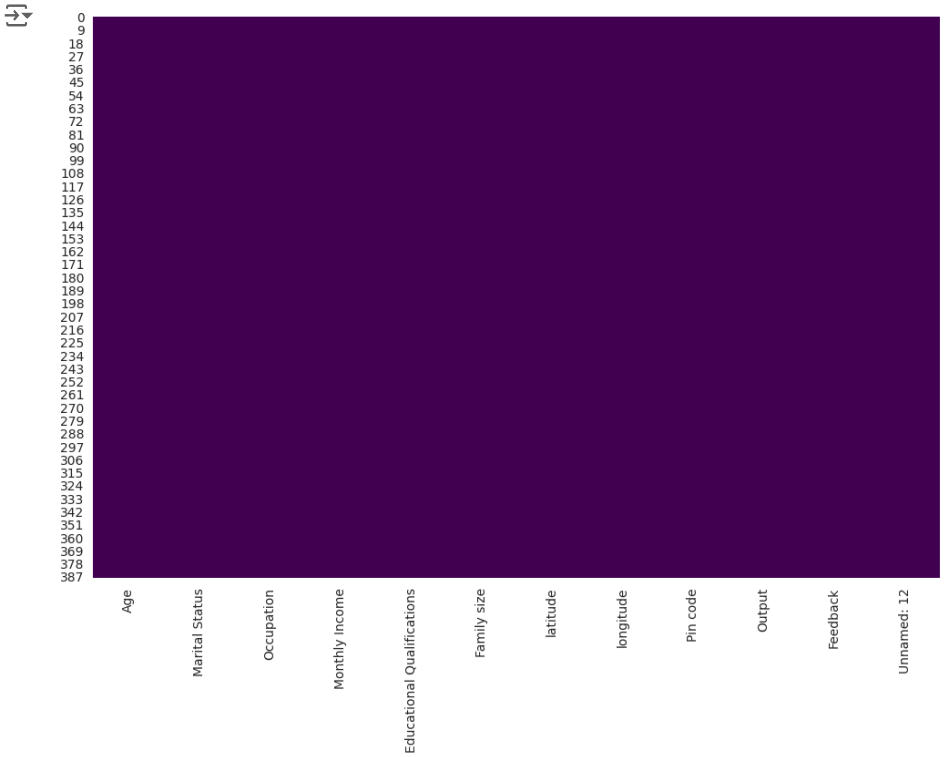

```
plt.figure(figsize=(10, 6))
sns.boxplot(data=df)
plt.show()
```



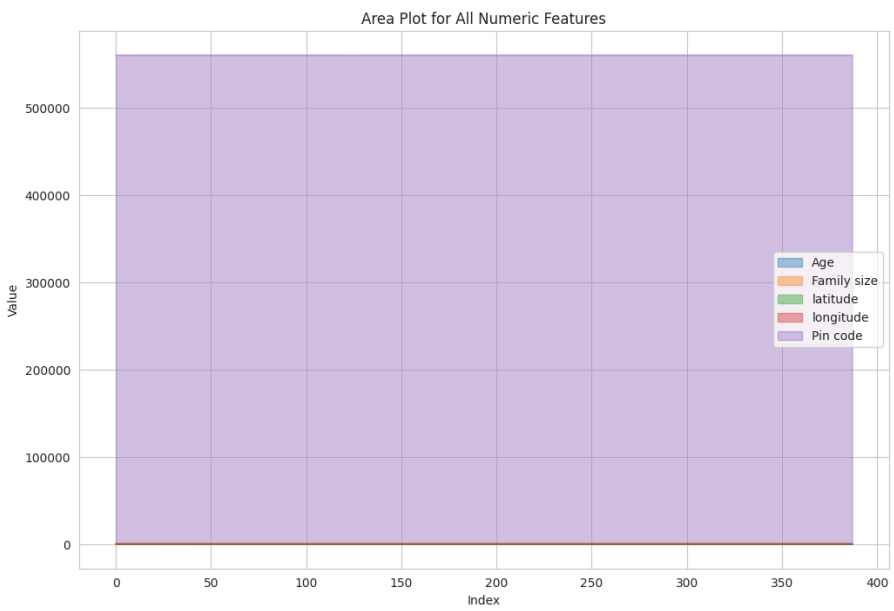

```
plt.figure(figsize=(10, 6))
sns.countplot(x='Feedback', data=df)
plt.show()
```




```
plt.figure(figsize=(12, 8))
sns.heatmap(df.isnull(), cbar=False, cmap='viridis')
plt.show()
```




```
df.plot(kind='area', alpha=0.4, figsize=(12, 8))
plt.title('Area Plot for All Numeric Features')
plt.xlabel('Index')
plt.ylabel('Value')
plt.show()
```




```
plt.figure(figsize=(10, 6))
sns.barplot(x='Occupation', y='Age', data=df)
plt.title('Bar Plot of Occupation vs Age')
plt.show()
```

```
plt.figure(figsize=(10, 6))
sns.violinplot(x='Occupation', y='Age', data=df)
plt.title('Violin Plot of Occupation vs Age')
plt.show()
```


```
plt.figure(figsize=(10, 6))
sns.countplot(x='Occupation', data=df)
plt.title('Count Plot of Occupation')
plt.show()
```